ERIC Document Resume	See Instruction Sheet	DATE OF RESUME	DATE RECEIVED

A. PUBLICATION DATA

1. ERIC NUMBER	2. ORIGINATOR CONTROL NO(S).	3. IDENTIFICATION NO.
ED002555	N.A.	2236

DO NOT WRITE IN SPACE BELOW

4. AUTHOR(S): Radlow, Robert

5. TITLE OF DOCUMENT ABSTRACTED:
The Relation of Some Measures of Ability to Measures of Learning from Sound Motion Pictures

6. PUBLICATION TITLE: N.A.

7. EDITOR(S): N.A.

8. PLACE OF ORIGIN
ORG. Instructional Film Research Program
ST. Pennsylvania State University
CITY, STATE State College, Pa.
TELE. EXT.
INST/LOC CODE

9. PUBLISHER OR SOURCE Instructional Film Research Program, Pennsylvania State University, State College, Pennsylvania

10. VOL.	11. PAGES	12. MO./YR.	13. TOTAL PP.
N.A.	N.A.	9/55	16

MICROFILM COPY AVAILABLE
YES ☒ NO ☐

COPYRIGHT GRANTED
YES ☐ NO ☒

COPYRIGHT RELEASE FILED
YES ☐ NO ☐

TYPE OF REPORT
PROPOSAL ☐
IN-PROGRESS ☐
FINAL ☒
FOLLOW-UP ☐

OTHER ERIC REFERENCES:

ERIC SATELLITE CODE

B. INDEXING DATA

14. SUBJECT AREA

15. TYPE OF ENDEAVOR Report

16. INDEXING TERMS (6-12)

audiovisual aids	sound motion pictures			
students	teaching methods	learning		
abilities	measurement	motion pictures		
achievement	instructional film			

17. ABSTRACT (MAX. 250 WORDS)

The relationship of various mental abilities to learning from films was studied to test whether instructional films should be prepared with audience characteristics clearly defined and whether films may benefit persons with certain mental abilities or deficiencies.

Measures of factual learning from two sound motion pictures were obtained for a group of basic trainees at Indiantown Gap Military Reservation. Various abilities were measured to get accurate descriptions of the subjects. A control group which did not see the films was given the same informational tests on the instructional films as the experimental group.

High verbal comprehension, general reasoning, and spatial orientation were characteristics of superior film learners. Perceptual speed was not a useful predictor of learning. The low reliability of scores in the study limited the magnitude of the correlations that could be obtained between those gains and the abilities measured. Verbal comprehension was not always the best predictor of film-learning; the ability which was the best predictor of film-learning depended upon the specific film and upon how learning from the film was measured. Several ability measures should have been used in selecting trainees.

ERIC No.

18. SOURCE OF ABSTRACT

C. FUNDING DATA

19. SUPPORTING ORGANIZATION	20. DIVISION OR DEPT.	21. CONTRACT NO. / PROJECT NO.
	22. TOTAL COST OF PROJECT	
	23. DURATION OF PROJECT	U.S. Department of Health, Education, and Welfare Office of Education

ERIC Form 1

2236

ED025555

TECHNICAL REPORT - SPECDEVCEN 269-7-58

THE RELATION OF SOME MEASURES OF ABILITY TO MEASURES OF LEARNING FROM SOUND MOTION PICTURES

Principal Investigator:

Robert Radlow

Dean M. R. Trabue
Responsible Administrator

C. R. Carpenter
Program Director

Pennsylvania State University
University Park, Pennsylvania

SPECDEVCEN Project 20-E-4
Contract N6onr-269

---

20 September 1955

For the Special Devices Center

Instructional Film Research
Distribution List

C. H. S. Murphy, Captain, USN
Commanding Officer and Director

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education

This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

3ND.PAPO

FOREWORD

Problem:

Instructional films should be prepared with the characteristics of the audience clearly defined. Films may be of particular value for persons with certain mental abilities or deficiencies. This research was carried out to discover the relationship of various mental abilities to learning from films.

Results:

High verbal comprehension, general reasoning and spatial orientation were found to be the characteristics of superior film learners.

Recommendation:

Training films should be used by trainees with high scores in verbal comprehension, general reasoning and spatial orientation.

LORAN C. TWYFORD, JR., Ph.D.
Staff Psychologist

CLIFFORD P. SEITZ, Ph.D.
Head, Human Engineering Department

THE RELATION OF SOME MEASURES OF ABILITY TO MEASURES OF LEARNING FROM SOUND MOTION PICTURES[1]

by

Robert Radlow

The Instructional Film Research Program
The Pennsylvania State University

INTRODUCTION

Knowledge about the relationship of various mental abilities to learning from sound motion films is of interest from both a relatively practical and a more theoretical point of view. The practical aspect includes the development of methods for the selection of superior film-learners as well as suggesting which films would be relatively more effective for a target audience which is already defined. Such a project can be of theoretical interest too if it could be shown that certain identifiable abilities are related to an ability to learn. It would be important to be able to identify any correlates of learning, for practically all studies purportedly dealing with learning (which by definition is the difference between what is known before and what is known after that experience) actually study only what is known after "learning." That is, most studies analyze achievement (what is known after some experience) rather than what is learned.

One important characteristic of the tests to be used in such a project is that the measures of ability be describable in some other manner than by producing the test items themselves: the test name should describe what is measured when used rather than what test constructors think it measures when they look at the items. In other words, unless superior methods are developed the ability tests must either be the product of a progressive factor-analysis (by which the factorial purity of the tests has been increased to the point at which it becomes possible to call them tests which rank people on some identified trait) or, if they are not factorially pure, they must then be factor-analyzed in order to know

---

1. This report is based on a thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at the Pennsylvania State University.

what has been measured. If this can be done, then it is possible to get information about the relation between learning in a defined task and the possession of certain describable abilities. The present study will present some evidence on whether it is possible to define a general film-learning ability in terms of well described variables related to film learning.

GENERAL PROCEDURE

Measures of factual learning from two sound motion pictures were obtained for a group of basic trainees in the Army. An attempt was made to get accurate descriptions of these individuals by the use of tests of five kinds of abilities. These measures were chosen to give as complete a description of the individuals studied as was possible in the time available for testing, and at the same time to include measures of abilities which the investigator thought to be important in learning from instructional films. It will soon become clear why the broadest possible description was an important consideration. The two films used differed from each other as much as possible in order to permit some generalization with respect to film learning.

A control group which did not see the films was given the same informational tests on the instructional films as the experimental group, so that it would be possible to know if the gain score used as a measure of learning was a result of seeing the instructional films.

The Films

Two instructional films were used. One of these, Introduction to Measuring Instruments, is a twenty-five minute black and white film which gives a simple introduction to the names and functions of a number of measuring devices commonly used in machine shop work. The vocabulary is appropriate to men with a twelfth grade education, and the material presented is quite easy to learn.

The second film, The Weather, is a fifteen minute film presenting a great deal of relatively complex information on meteorology, including some fairly abstract principles. The vocabulary in the film is appropriate to people who have a twelfth grade education in terms of the actual words used, but the concepts taught appear to be much more difficult to learn than the material presented in Introduction to Measuring Instruments, and the abilities required to learn from the films might be expected to be quite different for the two films.

One justification for this study was the idea that if there is some ability which is necessary to learning from films in general, a measure of this ability should be related to learning from both of these instructional films in spite of the systematic differences between them. The assumption is made that we can obtain ideas about the nature of film-learning by comparisons of what has been learned (by the same individuals) from films which have been selected to differ from each other in most respects. If that can be done, no resemblance between the learning tasks would create a pair of correlations that are spuriously close in magnitude. However, since it was not possible to investigate anything but the correlations of test scores themselves, there is then another source of a spurious resemblance in the magnitudes of correlations that may have nothing to do with the films: if learning is measured in both instances by the same kind of instrument, then a similarity is effectively introduced into the learning task. This is true because skill in understanding what questions in a test mean, and other skills related to good test performance, are not what is meant by learning; yet it is impossible to separate the true learning from the effects of such differences in test-taking skill. For example, one might employ two multiple-choice information tests with items framed in much the same manner. In that case there would be a systematic tendency for one group of people to obtain higher scores on both tests: those people would be people with greater skill in taking the particular kind of multiple-choice test as would be shown when by merely changing the form of the test the rankings of the individuals would be changed. To obtain information about the effect of employing different measuring instruments, two alternate tests covering the same items of information in different ways were used for one of these films. This subject will be discussed further.

Information Tests on the Films

The tests on the films were made up of multiple-choice questions with four alternatives which could be answered on an IBM answer sheet. The test on the film Introduction to Measuring Instruments had sixty such questions, which were primarily verbal, although some were based on drawings of the various measuring instruments. This test will be labeled MI in this report.

There were two tests used to measure gains resulting from seeing The Weather film. One of these was a sixty-four item multiple-choice test with four alternatives in which the items were entirely verbal. This

test is referred to as WA in this report. The first 34 items in this test were such that the stem or question part of the items could be used with drawings based on the film instead of verbal descriptions as they were in test WA. Therefore, a second test was made up composed of the 34 items in which the stem of each item included a drawing. This test is referred to as WB in this report.

The purpose of this second test was to determine whether abilities required to answer partly pictorial items were different from abilities required to answer entirely verbal items. Since only 34 of the 64 items in WA could be cast in the pictorial form, these 34 items in WA had to be scored separately to permit a direct comparison between verbal and pictorial items. The score on these 34 items in the verbal form is called WAS in this report, although it is not really a separate test at all.

Inasmuch as the purpose of this study was to investigate learning, it was necessary to administer these tests of knowledge twice. The first test was given before the film was seen to provide an index of knowledge of the subject before instruction. The second test was given immediately after the film showing to provide an index of what was known after instruction. The difference between these scores is an index of what has been learned. This difference score (or learning gain score) was then correlated with the various measures of ability (the factor-pure tests that have been mentioned and will be described below), and the correlation indicates how differences in these abilities are related to learning from these films. The more difficult question, and the more important one, is what relation if any can be inferred about film-learning in general. This question will be discussed below.

Measures of Related Abilities

The purpose of this project was to study the relationship between gains on these tests of information and measures of certain abilities. Several tests from the Guilford-Zimmerman Aptitude Survey were selected to give measures of abilities which might be related to learning from films. The survey includes tests which were developed on the basis of several factor-analyses. These successive factor-analyses were undertaken to determine the effect of variation of content upon factorial purity of the tests. Because of the method of factor-analysis employed this procedure tended at the same time to reduce the intercorrelations between the measures of abilities. Thus, successive improvements produced tests which not only measured identifiably different traits, but also were as nearly independent of each other as it was possible to make them. The use of such tests in this project seemed advisable, since it was hoped that the tests used might describe a wide range of abilities of the trainees

in the short time available for testing. Tests with low intercorrelations will be more efficient in providing information about abilities in the minimum testing time. Yet it should be clear from the introduction that factor-pure tests were desirable for a more important reason. No matter how many unfactored ability tests were employed in a correlational study of this sort, less information would be available about the abilities correlated with superior learning from films (unless a factor-analysis were performed on the tests after they were given). Names assigned to tests are always arbitrary in a sense. They may be names for what the test constructor hopes he is measuring or they may be names that indicate what many unbiased individuals believe the test is measuring. But more objective methods are necessary to verify that the instruments are at least principally measuring the trait they have identified with a name.

This problem is important enough to justify an example. Many people, feeling that they understood what tasks would separate good "reasoners" from poor ones, set about to publish them and called them "Reasoning Tests." Those with the most "common sense" presented situations in which "reasoning ability" was needed to reach a correct solution. To be sure, these people were often right: such tests would differentiate between people who could solve problems similar to the ones in the test and those unable to do so. However these tests typically demanded superior verbal or mathematical skill in order to achieve a good score. Thus, there would be cases in which a "good reasoner" would score poorly because he does not use words well, and there would also be "poor reasoners" who would score well because of superior verbal ability. It has been through factor-analysis that the true complexity of such tests that purport to measure a trait has been understood. To complete this example, it has been shown by the investigations of factor-analysts that specially constructed tests of arithmetic reasoning will measure the same reasoning ability that is demanded in the other reasoning tests, while at the same time a score on the test is less affected by abilities that are not included in the title of the test. This result may seem contrary to "common sense," but the finding is a reliable one.

If a conclusion is needed, it is that while imagination is good in the construction of tests, factor-analysis is an important check on intuition in deciding how to name what the test measures. Even so, a test must be named by considering what kind of tests measure the same or similar traits.

This example illustrates the importance of employing tests of proven simplicity which test identifiable traits. If the tests were unanalyzed, it would not be justifiable to assert that a trait is related to film-learning.

The tests selected from the Guilford-Zimmerman Aptitude Survey for use in this study were the following:

1. Verbal Comprehension (VC) which is a multiple-choice test of vocabulary definition.

2. General Reasoning (GR) which is a test of ability to solve problems that are primarily arithmetic reasoning problems. Ability to solve these problems seems to be independent of ability to do arithmetic computations.

3. Perceptual Speed (PS) in which the problem is to decide which of five pictures is identical to a given picture. It seems to require rapid recognition of small changes in detail.

4. Spatial Orientation (SO) which is a measure of ability to determine changes in the direction and position of a boat in two successive pictures from external cues.

5. Spatial Visualization (SV) which tests ability to imagine how an alarm clock would look if it were rotated about one or more of three perpendicular axes.

The Trainees

A single company of recruits at Indiantown Gap Military Reservation was divided into a control and an experimental group by platoons. It was reported that the men had been assigned to the platoons in a random manner. They were met on six different occasions for periods of time ranging up to two hours. During this time the tests were administered and the two training films were shown. The experimental and control groups had almost identical treatment except that the control group did not see the films. The testing sessions were arranged so that there would be a week between the pre-test on a particular film and the post-test (which was taken immediately after seeing the film). This delay was incorporated into the design of the study to minimize certain undesirable effects of taking the same test twice. For example, it was believed that a delay would reduce the boredom involved in test-taking. It would also

reduce the memory of specific items, which might serve as cues in fixating certain items of information from the film, or might merely sensitize the repetition of answers previously given (regardless of their correctness).

The schedule of testing for the two groups was as follows:

TESTING SCHEDULE FOR EXPERIMENTAL AND CONTROL GROUPS

Session	Experimental Group	Control Group
1	Pre-Test MI	Pre-Test MI
2	Pre-Test WB, WA	Pre-Test WB, WA
3 & 4	VC, GR, PS, SO, SV	VC, GR, PS, SO, SV
5	See film Introduction to Measuring Instruments Post-Test MI	Post-Test MI
6	See film The Weather Post-Test WB, WA	Post-Test WB, WA

RESULTS

Before the principal part of the analysis could be done for this study it was necessary to know that learning which was shown by the tests on Introduction to Measuring Instruments and The Weather resulted from having viewed the films and not from any other cause. Therefore, a control group was given these information tests at the same time as the experimental group was tested, both before and after the time the films were shown. But the control group did not see either film.

The data in Table 1 are presented as evidence that little or no learning gain took place when no film was seen.

TABLE 1

COMPARISON OF THE PRE- AND POST-TEST SCORES FOR THE CONTROL GROUP (N = 70)

	MI	WA	WB
Mean Gain	.78	-.20	.50
SE Mean Gain	3.93	5.15	3.18
CR	.20	-.04	.16
P	.84	.97	.88

In view of the small gains shown by the control group it may reasonably be assumed that any gain shown by members of the experimental group was due to having seen the film. The rest of the data on which this report is based were collected using the experimental group.

Table 2 is presented principally to show the extent to which the Guilford and Zimmerman tests were "factor pure" in this study. Ideally, intercorrelations approximating zero should have been obtained. The data show that this goal is far from attainment, but this does not controvert the fact that this was probably the best battery of achievement tests which was available for the purpose of this study.

TABLE 2

INTERCORRELATIONS OF ABILITY TESTS

	VC	GR	PS	SO	SV
VC	-	.41	.13	.33	.31
GR			.18	.31	.50
PS				.24	.37
SO					.53
SV					

(N is between 76 and 81)

In Table 3, the intercorrelations of the tests of ability and learning are presented, as well as the estimated reliabilities of the tests on the films.

The information in Table 3 is useful in identifying some of the abilities needed for learning from these two films. For example, the correlation of .53 between Verbal Comprehension and gain in MI indicates that verbal ability is relatively more important for learning from Introduction to Measuring Instruments than for learning from The Weather, since the correlations of Verbal Comprehension with tests on the latter film are considerably lower (.35, .23, .29). Perceptual Speed appears to be relatively unimportant for learning from either of these films, since the correlation of this test with any of the measures of learning are generally quite low. In fact, only two of the four correlations met the 5 per cent level criterion for a significant difference from zero. The correlation of Verbal Comprehension would be unlikely to surprise anyone, nor would some others that can easily be read in Table 3. But the virtual uselessness of Perceptual Speed in predicting film learning was unexpected. In this way the correlations presented in Table 3 provide an index of the value of each ability measure in predicting learning from each film.

TABLE 3

INTERCORRELATIONS OF ABILITY TESTS AND MEASURES OF LEARNING ON INFORMATION TESTS
(N is between 76 and 81)

	MI	WA	WB	WAS
VC	.53	.35	.23	.29
GR	.42	.31	.31	.24
PS	.11	.20	.12	.25
SO	.27	.28	.24	.31
SV	.45	.34	.20	.35
Est. Rel. Pre-Test*	.67	.35	.25	.25
Est. Rel. Post-Test*	.91	.82	.68	.68

* Reliability estimates computed by Kuder-Richardson Formula 14

Reading the table across rows shows whether a particular ability measure is an equally effective predictor for the four learning criteria. For example, the finding that VC correlates .53 with MI but only correlates .23 with WB indicates that verbal ability is less effective in predicting learning in the latter situation. However, more comparable comparisons can be made between WAS and WB because they are two different types of tests of the same information and have equal reliabilities. There are differences here, too. The importance of the differences between the WB (verbal-diagram test) and WAS (all verbal test) columns will be discussed in the next section.

Another purpose of this study was to develop useful measures for efficient description of a target audience that is to be taught by sound motion films. The test weights for predicting learning from these films are presented in Table 4, together with the multiple correlation coefficients. If a group of individuals were given these ability tests and their scores were multiplied by the appropriate regression weights, then the sum of these products would represent the best prediction (from these data) of which individual would learn the most from that film. The higher the individual's total score, the higher the predicted learning score would be.

TABLE 4

REGRESSION WEIGHTS FOR PREDICTING LEARNING FROM TWO INSTRUCTIONAL FILMS

	MI	WA	WB	WAS
VC	.41	.23	.09	.17
GR	.12	.11	.24	.02
PS	-.07	.08	.05	.13
SO	-.05	.08	.15	.12
SV	.32	.15	-.04	.17
$R_{-.12345}$	.62	.45	.37	.43

The multiple correlation given at the bottom of Table 4 is the correlation between the learning scores these subjects actually obtained and the learning scores this method would predict for them. All of these multiple correlations show that employing the tests together resulted in substantial improvement in prediction over the accuracy obtainable with only one test.

But it must be noted that the absolute magnitude of these correlations may be misleading because the gain scores were rather unreliable, and this unreliability reduces the magnitude of the correlation and multiple correlations with learning reported in this study. The problem of the unreliability of gain scores will be discussed in the next section.

DISCUSSION

Few people would be interested solely in predicting learning resulting from seeing these particular films. They would be more likely to express interest in predicting which people would learn most in an established training program that relied heavily upon instructional films. It is clear now that this is possible. It is also clear that it would require the administration of an extensive battery of tests and the use of different regression weights for each film. Thus, it would be of greater utility and interest to be able to predict learning by the use of one set of tests and regression weights that would be useful for all films or for broad classes of films.

The evidence is quite clear that Perceptual Speed is not a useful predictor of learning from films such as these. Perhaps this is true because neither film has a very rapid development. The constructive side concerns which abilities besides Verbal Comprehension would be most useful in selecting the superior film-learners. The issue is so stated because ability tests, at least highly loaded in the verbal factor, are regularly employed in selection where any selection program exists. The important question is: What else is important? The best answer on the basis of these data is that General Reasoning and Spatial Orientation (if not other tests as well) should be used in selection. In fact it is likely that these tests will become relatively better predictors of learning when applied to populations which have been "picked over" on the criterion of verbal proficiency. (Note that this process had not yet begun with the subjects in this study.)[1]

It may also be worthy of mention that some evidence exists that Spatial Orientation is an ability which is likely to be moderately important in predicting learning from any film. The evidence is not clear enough

---

1. What this means is that if the variability of a given predictor is systematically reduced, that predictor will become less effective. Thus, when brighter people are selected for special training, they are usually differentiated by tests of verbal skill. In consequence tests of verbal skill will be of less value in predicting other performances since their use for the first purpose has reduced the variability of verbal skill in that population more than the variability of other skills.

to state that it is a measure of film-learning ability, or even that a modification of that test would result in an index of film-learning ability. However, the correlations of Spatial Orientation with the four learning scores are much more constant than the correlations of any of the other four measures of ability. (This is even apparent in Table 3 by noting the range of variation of the correlations of learning scores with each of the ability measures. Statistical comparisons can be found in the thesis on which this report is based.) This relation would be important if it could be shown to hold for other films. Thus it is possible that a film-learning trait exists although Woodrow[1] has presented some evidence that no general learning ability exists.

It was noted earlier that VC correlates .53 with MI but only correlates .23 with WB. This is an interesting comparison because it suggests that verbal ability is relatively more important in learning a more highly verbal task. However, this inference could not be drawn from the study because of other differences in the tasks to be learned, and because of many differences in these measures of learning.

Comparisons between WB and WAS are of interest because they are tests of learning the same material and have almost identical reliabilities. Such comparisons can show that measuring the same learning with instruments of different forms results in different relations between the amounts learned and levels of ability. In this situation a diagram test and a verbal test on the same information showed different amounts of learning. This study was not designed to determine what relation there is between the possession of certain special abilities and differences in learning when only one identifiable characteric of a film is changed. But this study shows that if such an interesting study were undertaken, it would be an important and difficult task to select unbiased measures of learning.[2]

In considering the results of this study, it should be emphasized that the reliability of measures of gain are often quite low, even though the reliability of the test from which the measure of gain is derived are fairly high.[3] Since the measure of gain is based on two observations of the

---

1. Woodrow, H., "Factors in Improvement with Practise," Journal of Psychology, 7, 1939, 55-70.

2. Such a study has since been undertaken. See Lefkowith, E.F., The Validity of Pictorial Tests and Their Interaction with Audio-Visual Teaching Methods," Technical Report 269-7-49, Penn State University.

3. See Watkins, R. W., "An Investigation of the Reliability of Measures of Gain," unpublished doctoral thesis, Syracuse University, Syracuse, New York, August 1955.

same individual's behavior, errors of observation may be made twice, so that the errors of observation of gain scores is often large. The reliabilities of the gain scores used in this study were not estimated directly, but they may well be as low as .35 or .40 for the four measures of gain used. This estimate is based on indirect methods which have been attempted.

The principal effect of the low reliability of scores in this study is to limit the magnitude of the correlations that could be obtained between those gains and the abilities measured. Since low reliabilities suggest that a large part of any observation may be error which is not related to any variable, it is not possible to find variables which will have high correlation with the unreliable measure. While there are some statistical techniques available which would give estimates of the correlations that could be obtained if more reliable measures had been used, this procedure did not seem justified in the absence of really accurate data on reliability.

IMPLICATIONS

The study reported here was undertaken principally to discover how superior film-learners could be best identified and selected. It has been shown that Verbal Comprehension may not always be the best predictor of film-learning, but that the ability which is the best predictor of film-learning depends upon the particular film and upon how learning from the film is measured. Thus it is desirable to employ several ability measures whether the purpose is to classify or select individuals in training, or to secure data for matching purposes in experimental work. Several considerations suggest that Verbal Comprehension, General Reasoning, and Spatial Orientation should provide a satisfactory minimal battery for these purposes. It must also be suggested, in view of the results, that the validity of alternate criteria for measuring learning always be investigated. This is important because there is no way to distinguish between the "amount learned" and how the learning is measured.

Spatial Orientation is of special interest for two reasons. First, when selection of trainees must be made from populations which have already been selected on the verbal fluency variable, it is likely that this ability will be of relatively greater utility in predicting learning. Furthermore, the relatively small fluctuations between the correlations of Spatial Orientation with the different criteria of learning suggest that it might be a particularly useful predictor of film learners. It should be possible to obtain trainees who learned (and performed) the criterion task well and yet were inferior in these other abilities merely by employing small selection ratios. This practice could lead to a more efficient

use of manpower, and at the same time might improve understanding of what the term "film-learning ability" means.

Research on this project has lead to the realization that reliable data on learning (as opposed to post-experience measurements only) is almost non-existent. This points to the need for intensive research on the subject, since "learning" is so often the criterion in psychological research.[1]

---

1. This problem has recently been the subject of study by Richard W. Watkins of the Instructional Film Research Program. The study is available from the Syracuse University, New York, in the form of a doctoral dissertation entitled, "An Investigation of the Reliability of Measures of Gain."